



DFG

# Datenkorpora für Künstliche Intelligenz

DFG-Ausschreibung vom 3. April 2025

*Hinweis: Es gelten im Zweifel die Regelungen in der [Ausschreibung](#)*

Inhaltlich: Dr. Stefanie Mewes, Dr. Matthias Katerbow, formal: Clara Grau

# Ablauf des Webinars

1. Eckpunkte der Ausschreibung
2. Antragsvorbereitung – was ist zu beachten?
3. Ihre Fragen & Antworten

Die Folien werden auf der Website der Ausschreibung veröffentlicht

# Ziel der Ausschreibung

- ▶ **Auf- und Ausbau** qualitativ hochwertiger, umfangreicher **Datenkorpora**
  - Datenbasis für die (Weiter-)Entwicklung und Anwendung von KI-Methoden für die Wissenschaft
  - Einsatzbereich der KI-Methoden: **Forschung & wissenschaftliche Informationsinfrastrukturen**
- ▶ **Bedarfsorientierung und Einsatzpotential**
  - Qualität, Umfang und Zusammensetzung der Datenkorpora an Bedarfe abgestimmt
  - über einzelne Forschungsfragen und Standorte hinweg Forschung ermöglichen oder die Informationsversorgung verbessern
- ▶ **Bereitstellung der Datenkorpora**
  - bewährte Prinzipien (FAIR, CARE) und **Standards berücksichtigen**
  - über vorhandene Informationsinfrastrukturen, insbes. NFDI, erfolgen



## KI – *breites Methodenspektrum*

- Maschinelles Lernen
- Text- und Data Mining
- ...

- ▶ Auf- und Ausbau von Datenkorpora für Künstliche Intelligenz
  - Konzeption von Auswahl- und Qualitätskriterien sowie Umsetzung von Qualitätssicherungsmaßnahmen
  - Nachnutzung und gegebenenfalls Anpassung sowie insbesondere Anwendung von Verfahren der Datenbereinigung, -aggregation, -annotation bzw. -kuration und -harmonisierung
  - Verfügbarmachung und Bereitstellung des aufzubauenden Datenkorpus innerhalb bestehender wissenschaftlicher Informationsinfrastrukturen



# Antragstellung

Deutsche Forschungsgemeinschaft  
Informationen für die Wissenschaft

## Ausschreibung

Nr. 28 3. April 2025

### Datenkorpora für Künstliche Intelligenz

DFG fördert den Aufbau von Datenkorpora zum Trainieren von Künstlicher Intelligenz (KI).

Der Ausschuss für Wissenschaftliche Bibliotheken und Informationssysteme (AWBI) der Deutschen Forschungsgemeinschaft (DFG) greift Bedürfnisse der wissenschaftlichen Community auf, die im Rahmen des [Leitfadens zur Unterstützung von KI in der Forschung durch Informationsinfrastruktur](#) formuliert wurden. Mit der vorliegenden Ausschreibung können Projekte gefördert werden, in denen die Aufbereitung und zur Verfügungstellung von Daten als Grundlage für die Weiterentwicklung von KI für die Wissenschaft erfolgt.

#### Hintergrund

Methoden, die im Kontext der Künstlichen Intelligenz angewendet werden, wie maschinelles Lernen (ML) und Text- und Dokument (TDM), haben in weiteren Forschungsbereichen sowie die wissenschaftlichen Informationsversorgung auf neue Weise zur Analyse und Verbindung großer Datenmengen sowie zur Speicherung, Verwaltung und Anwendung dieser Methoden beigetragen. Die Entwicklung und Anwendung dieser Methoden besteht aus verschiedenen Daten (Sachdaten, Metadaten, Vermerksdaten, ein- und mehrsprachige, strukturierte Daten, Digitalisate, Profillängendaten etc.), die in ihrer Komplexität teilweise erhebliche Verfahren und deren Aggregation und Bereitstellung beträchtliche Aufwände verursachen. Demzufolge bedarf es in der Wissenschaft systematisch in Bibliotheken, Kuratieren, annotierten und aggregierten Daten.

#### Ziele der Ausschreibung

Die Förderung soll den Aufbau und Ausbau qualitativ hochwertiger, umfangreicher Datenkorpora ermöglichen, um eine breite und wissenschaftlich fundierte Basis für die Weiterentwicklung und Anwendung von Methoden der Künstlichen Intelligenz für die Wissenschaft zu schaffen. Der künftige Einsatz von Methoden und Anwendungen, die auf den geänderten Datenkorpora beruhen werden, kann dabei sowohl in der Forschung als auch innerhalb wissenschaftlicher Informationsinfrastrukturen liegen. Qualität, Umfang und Zielausgangspunkt der Datenkorpora müssen dabei auf die jeweiligen Bedarfe abgestimmt sein und unter der Prämisse stehen, über einzelne Forschungsfragen und Standorte hinweg Forschung zu ermöglichen oder die Forschungsversorgung zu verbessern. Die Bereitstellung der Datenkorpora soll sowohl in Prinzipien (FAIR, CARE) und Standards berücksichtigen und über vorhandene Informationsinfrastrukturen, insbesondere über Nationalen Forschungsdatenzentrum (NFZ), erfolgen.

Deutsche Forschungsgemeinschaft  
Kennedyallee 40 · 53175 Bonn  
Telefon: +49 228 885-2109 · Telefax: +49 228 885-2180 · presse@dfg.de



Ausschreibung

DFG-Vordruck 12.14 – 04/22 Seite 1 von 13

## Merkblatt und ergänzender Leitfaden

### Informationsinfrastrukturen für Forschungsdaten

Deutsche Forschungsgemeinschaft  
Kennedyallee 40 · 53175 Bonn  
Telefon: +49 228 885-1 · Telefax: +49 228 885-2777 · postmaster@dfg.de · www.dfg.de



DFG-Vordruck 12.14

DFG-Vordruck 12.01 – 03/25 Seite 1 von 15

## Leitfaden für die Antragstellung

### Projektanträge im Bereich Wissenschaftliche Literaturversorgungs- und Informationssysteme (LIS)

Deutsche Forschungsgemeinschaft  
Kennedyallee 40 · 53175 Bonn  
Telefon: +49 228 885-1 · Telefax: +49 228 885-2777 · postmaster@dfg.de · www.dfg.de



DFG-Vordruck 12.01

Spezifische Regelungen ersetzen  
allgemeine

# Anforderungen zur Erfüllung der Förderbedingungen

- ▶ **geeignete Lizenz** zur kostenlosen Nutzung in der Wissenschaft
- ▶ **Verfügbarkeit** und **Zugang** zum künftigen Datenkorpus ist für wissenschaftliche Nutzer\*innen so **offen wie möglich** zu gestalten [...]
- ▶ Signifikante Teilergebnisse nach dem **ersten Projektjahr veröffentlichen**
- ▶ Datenkorpus in fachlichen und überfachlichen Verzeichnissen, Registries oder ähnlichem, **auffindbar machen**
- ▶ **Dokumentation** des Datenkorpus **gemäß einschlägiger Qualitätsstandards**
- ▶ **Doppelförderung ausgeschlossen**

# Antragsberechtigung

- ▶ Siehe [DFG-Vordruck 12.14](#) Ziffer 2.1
- Antragsberechtigt sind grundsätzlich **Angehörige von wissenschaftlichen Informationsinfrastruktureinrichtungen** wie bspw. Bibliotheken, Archiven, Museen, Forschungssammlungen, Forschungsdatenzentren oder Rechen- und Informationszentren u. ä., sofern sie gemeinnützig sind. Ferner ist **jede\*r Wissenschaftler\*in in der Bundesrepublik Deutschland** oder an einer deutschen Forschungseinrichtung im Ausland antragsberechtigt, dessen oder deren Ausbildung – in der Regel mit der Promotion – abgeschlossen ist.
- In der Regel nicht antragsberechtigt sind Sie, wenn Sie in einer Einrichtung arbeiten, die nicht gemeinnützig ist, oder Ihnen die sofortige Veröffentlichung der Ergebnisse in allgemein zugänglicher Form nicht gestattet ist.
- Da die Förderung im Bereich „Wissenschaftliche Literaturversorgungs- und Informationssysteme“ auf eine überregionale Verbesserung der Informationsinfrastrukturen abzielt und die daraus resultierenden Ergebnisse eine Dienstleistung für die Wissenschaft insgesamt darstellen, sind **Mitarbeiter\*innen der Institute und Mitgliedseinrichtungen der Max-Planck-Gesellschaft, der Fraunhofer-Gesellschaft, der Helmholtz-Gemeinschaft** oder der **Wissenschaftsgemeinschaft Gottfried Wilhelm Leibniz** sowie Angehörige von mit diesen Organisationen assoziierten Forschungseinrichtungen, die aus öffentlichen Mitteln grundfinanziert werden, und Angehörige deutscher Standorte international getragener Informationsinfrastruktureinrichtungen ebenfalls antragsberechtigt.

# Beantragbare Mittel und Projektdauer

## ▶ Obergrenzen

- 400.000 EUR
- 2 Jahre

## ▶ Beantragbare Module

- Basismodul

Personal- und Sachmittel - *ACHTUNG: Keine Investitionsmittel in dieser Ausschreibung!*

- Projekt-spezifische Workshops

Zuschuss zu den Gesamtkosten des Workshops, z.B. Fahrt- und Aufenthaltskosten der TN, Kosten für Organisation

# Fristen



Unverbindliche **Absichtserklärung**

28. Mai 2025



elan-Portal Registrierung

23. Juli 2025



**Antrag** via [elan-Portal](#)

**30. Juli 2025**

# Absichtserklärung

## ► Welche Angaben werden abgefragt?

<b>Antragsteller*in</b>	Name
	Institut, Einrichtung
	Email-Adresse
Weitere Antragsteller*innen	Name
	Institut, Einrichtung
Beteiligte	Name
	Institut, Einrichtung
<b>Vorhaben</b>	Titel
	Fach nach DFG-Fachsystematik
	Mittelvolumen
Kommentar	

# Begutachtung & Entscheidung

- ▶ Fachliche und informationsfachliche Aspekte
  - ▶ Kriterien gemäß der Ausschreibung
  - ▶ Allgemeine Hinweise für die Begutachtung ([DFG-Form 10.20](#))
  - ▶ Hinweise für die Begutachtung im Bereich LIS ([DFG-Form 10.213](#))
- 
- ▶ Anfang 2026 Kommunikation der Entscheidungen
  - ▶ Auftaktworkshop
    - Einladung an Projektnehmer\*innen
    - Weitere Informationen folgen



# Ablauf des Webinars

1. Eckpunkte der Ausschreibung
2. Antragstellung – was ist zu beachten?
- 3. Ihre Fragen & Antworten**



Bitte nutzen Sie die Funktion „Hand heben“



# Vielen Dank für Ihre Aufmerksamkeit

## Weitere Informationen

- ▶ zur DFG: <https://www.dfg.de>
- ▶ zum Förderatlas: <https://www.dfg.de/foerderatlas>
- ▶ zu allen geförderten Projekten: <https://www.dfg.de/gepris>
- ▶ zu den deutschen Forschungseinrichtungen: <https://gerit.org>
- ▶ zu historischen DFG-Projekten: <https://gepris-historisch.dfg.de>